Eur. J. Clin. Chem. Clin. Biochem.
Vol. 29, 1991, pp. 237–240

The Development of Reference Methods in Clinical Chemistry[1])

The contribution of the Community Bureau of Reference of the Commission of the European Communities

By *Elisabeth Colinet*

Commission of the European Communities (CEC), Community Bureau of Reference (BCR), Brussels, Belgium

(Received January 22, 1991)

Summary: The use of reference method procedures in clinical chemical analysis is advocated by many experts as the most reliable approach to obtaining accurate results. The performance of such procedures must, however, be rigorous. This contribution will emphasize the importance of interlaboratory studies for this purpose. Examples will be presented, taken from the work done under the BCR programme of the Commission of the European Communities. The determination of steroid hormones in serum by isotope dilution mass spectrometry and the measurement of enzyme catalytic activities, according to IFCC recommended methods, will be discussed.

The Concept of Accuracy

One of the often quoted definitions for a reference method is: "a method which after exhaustive investigation has been shown to have negligible inaccuracy" (1a). *Büttner* (1b) earlier pointed out the imortance of assessing accuracy. This contribution is intended to draw attention to the difficulty of demonstrating accuracy.

Before getting into complex measurement processes, such as those involved in the field of clinical chemistry, it seems appropriate to start with an example taken from the metrology field. At the top of the measurement hierarchy stand the primary Metrology Institutes. Their role is to provide the means of calibration for all base and derived quantities. Metrology institutes have developed quite an expertise in the identification and quantification of measurement errors. The most difficult task however is still the appraisal of the systematic component of the measurement error.

The problem can be illustrated by the results of an intercomparison of mass measurements organised by the Community Bureau of Reference and involving European Metrology Institutes (2). The purpose was to measure a 50 g mass. The results were to be expressed as corrections to the nominal value of 50 g. The results are shown in figure 1 together with the calculated uncertainties (as vertical bars). It should be noted that in Metrology laboratories, the uncertainties are derived from a combination of the variances of both random and systematic errors. Looking at these results and in particular at the scale of the plot, one first realizes that the between-laboratory differences are extremely small. However one participant (Lab 6) evidently had a problem that he did not suspect, as the uncertainty of the measurement results was underestimated. This suggests that even for a base quantity such as a mass, and for well trained laboratories the claim of absence of inaccuracy is a risky statement. If all the participating laboratories were able to assess their inaccuracy, all the uncertainty intervals as established here would overlap each other. In fact the intercomparison enabled the discrepancy to be perceived by Lab 6. This and other examples show that the search for the "true" value of a quantity should preferably result from a comparison process.

[1]) Based on a lecture given at the Symposium "Reference Methods in Clinical Chemistry – Objectives, Trends, Problems" of the Congress Biochemische Analytik 90, München, May 8, 1990

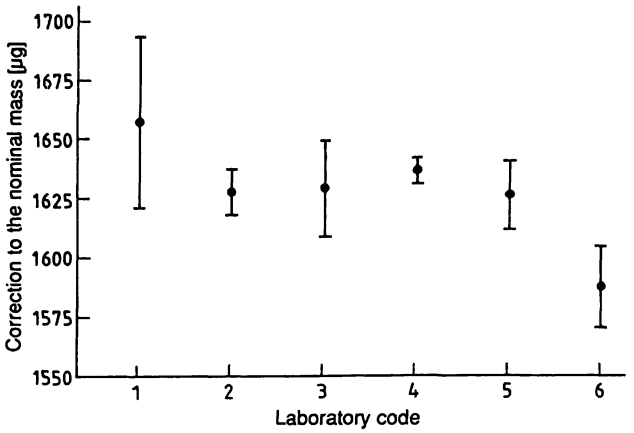

Fig. 1. Results of an intercomparison for the measurement of a 50 g mass, involving six Metrology Institutes. Results are expressed as deviations from the nominal 50 g value. Bars represent the uncertainties calculated by the participants, including random and systematic components.

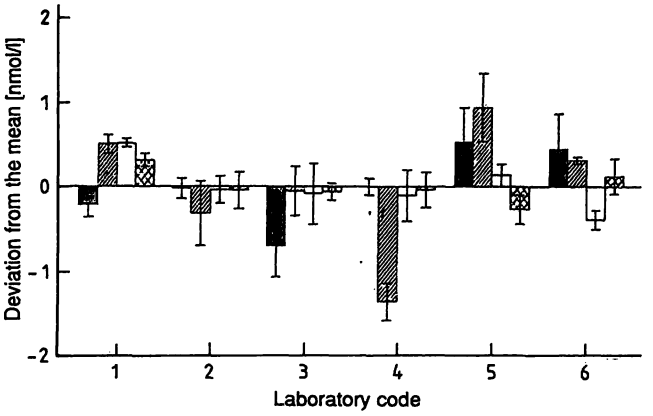

Fig. 2. Results of a series of four intercomparisons for the measurement of serum progesterone, involving six laboratories using an isotope dilution/mass spectrometry procedure. On each sample, and for each laboratory, results are expressed as the deviation of the laboratory mean relative to the mean of all laboratories. In each case, the error bar represents 2 SD of the laboratory mean.

1985 lyophilized
1986 liquid
1988 liquid
1989 lyophilized

Let us now consider the field of clinical chemistry where considerable effort has been devoted to the development of reference methods. One measurement principle currently presented as a sound basis for a reference method in many applications is that of isotope dilution mass spectrometry (IDMS) (3). The accuracy achievable by this measurement principle must, however, be established, and this can only be achieved in a reliable manner by interlaboratory studies.

Selected Examples

To exemplify the need for collaborative effort, the results of a series of interlaboratory studies conducted by the Community Bureau of Reference will be shown.

The work done for the determination of progesterone in serum is taken as the first example. Six European laboratories, all using an isotope dilution/mass spectrometry (IDMS) procedure, were involved. It is interesting to show the progressive yet difficult improvement achieved by this group of laboratories. Figure 2 shows the deviation of each laboratory mean from the mean of the means of four consecutive trials. The error bar stands for the random component of the error i.e. 2 standard deviations of the laboratory mean in each case. Each time, a different serum sample was analysed to assess to specificity of the procedure under test in each laboratory. The progesterone concentration was similar in each of the four serum samples and of the order of 10 nmol/l. Going from the lyophilized serum sample assayed on the first occasion to the liquid serum examined next, the between-laboratory variation increased. This worsening led the participants to scrutinize their procedures. The consequence of that investigation was that on the following liquid sample the agreement between the results was noticeably better. Figure 3 presents the overall improvement over the four year period in terms of both between- and within-laboratory coefficients of variation. Maybe it could be argued that all the participants had not sufficiently validated their procedures before starting. On the other hand, each of them could testify that it is only the confrontation of the results that permitted unsuspected sources of errors to be identified.

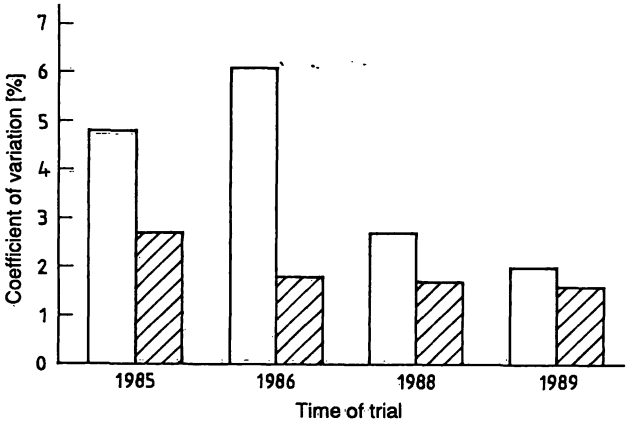

Fig. 3. Results of four intercomparisons for the measurement of serum progesterone, involving six laboratories using an isotope dilution/mass spectrometry procedure. Results are expressed in terms of both between- and within-laboratory coefficients of variation.

Between laboratory
Within laboratory

There comes a stage where no further improvement can be made. This should be taken as an estimate of the analytical state-of-the-art. Figure 4 shows the results obtained by the six laboratories on the last lyophilized serum sample. The vertical bar again represents 2 standard deviations of the laboratory mean.

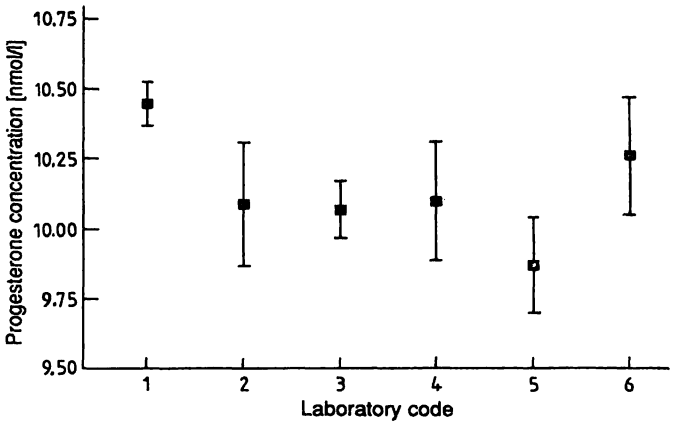

Fig. 4. Results of the certification exercise for the measurement of serum progesterone in CRM 347, involving six laboratories using an isotope dilution/mass spectrometry procedure. For each laboratory, the results are expressed as the laboratory mean value ± 1 SD.

It was suggested that such a realization of the measured quantity should be maintained and made available to help transfer the reference method value. Therefore the Community Bureau of Reference (BCR) was asked to issue Certified Reference Materials (CRM) with values assigned by laboratories that had analysed them until it became impossible to further resolve their discrepancies. The last lyophilized serum sample was therefore established as CRM 347.

Currently available BCR CRMs for steroid hormone determination in human serum are:

CRM 192 and CRM 193
for cortisol in human serum

CRM 347 and CRM 348
for progesterone in human serum,

each corresponding to a different concentration of the hormone (4, 5).

This first example presented the case of a reference method enabling in principle an assignment of an unbiased estimate of the "true" value of a well defined quantity in terms of the amount-of-substance concentration. The second example will address the problem of the determination of more complex parameters such as the enzymes. The approach first requires the selection of the means of expressing the value of the parameter. In the case of enzymes it is the catalytic activity concentration.

The International Federation of Clinical Chemistry (IFCC) has made a great contribution to the establishment of recommended methods, in particular in the field of enzymology. To date, to the best of our knowledge, three methods have been approved, namely those for alanine and aspartate aminotransferase and for creatine kinase. Provisional recommendations have been issued with regard to methods for γ-glutamyltransferase and alkaline phosphatase. In each case, the measurement procedure has been carefully assessed with regard to possible interferences by the components of the measurement system. However, it still appears important to assess the transferability of the method.

Again, the Community Bureau of Reference organises interlaboratory studies for this purpose. In particular, the measurement of alkaline phosphatase was examined for which a provisional recommendation had already been issued in 1983 (6). The results obtained for alkaline phosphatase determination in two successive trials are shown in table 1. The number of laboratories involved in both trials was fourteen. It must be noted that the samples submitted to the investigation consisted here of portions of a partially purified enzyme preparation. This choice was dictated by the wish to work with a well characterised material, the method having been previously assessed on serum samples. Because of the well known problem of reactivation of lyophilized alkaline phosphatase on rehydration, the first sample submitted to the interlaboratory trial was a liquid material. The IFCC-recommended method was used by all participants. However, some of them admitted that they were not familiar with it and this resulted in a wider scatter between laboratories than expected. Another trial was organised to test, this time, a lyophilized sample for which a defined protocol for reconstitution was laid down. In addition, each laboratory carried out further verifications to ensure the proper functioning of the instrumentation used. The result was a net improvement of the consistency of the laboratory sets.

Tab. 1. Determination of the catalytic concentration of alkaline phosphatase in two consecutive trials, each involving a different preparation of the partially purified enzyme. The number of participants was fourteen in each case. They all followed the IFCC recommended method.

	Mean concentration [μkat/l]	Between laboratory SD [μkat/l]	Between laboratory CV [%]
1st trial	7.69	0.75	9.8
2nd trial	4.23	0.18	4.3

Figure 5 shows, for each laboratory, the mean value with 2 standard deviations, obtained in the first exercise. Note the considerable within-laboratory variation for Lab 11, which had serious problems obtaining reproducible results from day-to-day, and the large bias of Lab 8. Figure 6 shows the data of the second exercise, where similar performance was achieved by all participants with regard to within-laboratory variation. Also observe the scale interval, which was reduced to 1 μkat/l.

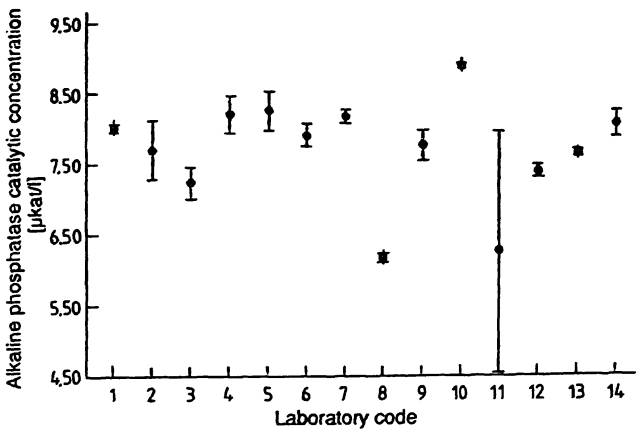

Fig. 5. Results of an interlaboratory study for the measurement of alkaline phosphatase catalytic activity, involving fourteen laboratories using the IFCC method. For each laboratory, results are expressed as the laboratory mean value ± 1 SD.

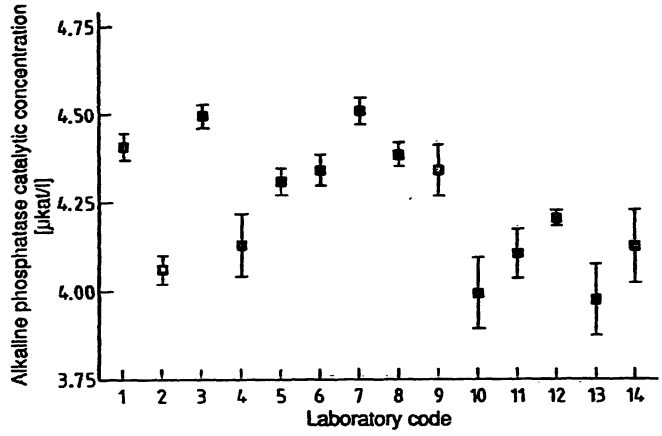

Fig. 6. Results of the certification exercise for the measurement of alkaline phosphatase catalytic activity in CRM 371, involving fourteen laboratories using the IFCC method. For each laboratory, results are expressed as the laboratory mean value ± 1 SD.

As in the case of the serum materials analysed for steorid hormones, enzyme materials were deemed useful as transfer standards. Hence the lyophilized preparation of alkaline phosphatase assayed in the second exercise became CRM 371. The following have been certified by the BCR:

CRM 319 for γ-glutamyltransferase (7)

CRM 371 for alkaline phosphatase (8)

Conclusions and Future Trends

The development of reference methods is generally felt to be indispensable, in order to generate reproducible and reliable results. It is increasingly acknowledged that reference methods, associated whenever necessary to reference materials, should become an essential component of a Quality Assurance System.

Therefore the Commission through its Community Bureau of Reference will support all endeavours to develop reference method procedures, and to contribute to their validation. The collaboration established in the fields of hormonology and enzymology, which proved to be both efficient and fruitful, is being strengthened and will be enlarged to embrace other branches of clinical chemistry and possibly other areas of the biomedical analysis.

References

1a. Büttner, J., Borth, R., Boutwell, J. H., Broughton, P. M. G. & Bowyer, R. C. (1979) Approved Recommendation (1978) on Quality Control in Clinical Chemistry, Part 1. General Principles and Terminology. Clin. Chim. Acta *98*, F129–F143; J. Clin. Chem. Clin. Biochem. *18*, 69–77 (1980).

1b. Büttner, J. (1991) Reference methods as a basis for accurate measuring systems. Eur. J. Clin. Chem. Clin. Biochem. *29*, 223–235.

2. Commission of the European Communities, BCR Information – Applied metrology (1989) Intercomparison of Mass Measurements. Report EUR 12234 EN.

3. De Leenheer, A. P., Lefevere, M. F., Lambert, W. E. & Colinet, E. S. (1985) Isotope Dilution Mass Spectrometry in Clinical Chemistry. Adv. Clin. Chem. *24*, 111–161.

4. Commission of the European Communitties, BCR Information – Reference materials (1985) The Certification of Cortisol in two lyophilized serum materials CRM 192 and CRM 193. Report EUR 9661 EN.

5. Commission of the European Communities, BCR Information – Reference materials (1989) The Certification of progesterone in two lyophilized serum materials CRM 347 and CRM 348. Report EUR 12282 EN.

6. International Federation of Clinical Chemistry. Expert Panel on Enzymes. IFCC methods for the measurement of catalytic concentration of enzymes, IFCC document with a view to an IFCC recommendation. Part 5. IFCC method for alkaline phosphatase (1983) J. Clin. Chem. Clin. Biochem. *21*, 731–748.

7. Schiele, F., Muller, J., Colinet, E. & Siest, G. (1987) Production and certification of an enzyme reference material for gamma glutamyltransferase (CRM 319). Part 1 and Part 2. Clin. Chem. *33*, 1971–1982.

8. Schiele, F., Muller, J., Colinet, E. & Siest, G. Certification of an enzyme reference material for alkaline phosphatase (CRM 371). Clin. Biochem. (to appear).

Dr. Elisabeth Colinet
CEC
BCR
200 Rue de la Loi
B-1049 Brussels